

Tilburg University

Statistics of Heteroscedastic Extremes

Einmahl, J.H.J.; de Haan, L.F.M.; Zhou, C.

Publication date:
2014

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Einmahl, J. H. J., de Haan, L. F. M., & Zhou, C. (2014). *Statistics of Heteroscedastic Extremes*. (CentER Discussion Paper; Vol. 2014-015). Econometrics.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

No. 2014-015

STATISTICS OF HETEROSCEDASTIC EXTREMES

By

John H.J. Einmahl, Laurens de Haan,
Chen Zhou

18 February, 2014

ISSN 0924-7815
ISSN 2213-9532

Statistics of heteroscedastic extremes

John H.J. Einmahl

Laurens de Haan

Tilburg University

Erasmus University Rotterdam and University of Lisbon

Chen Zhou

De Nederlandsche Bank and Erasmus University Rotterdam

February 18, 2014

Abstract. We extend classical extreme value theory to non-identically distributed observations. When the distribution tails are proportional much of extreme value statistics remains valid. The proportionality function for the tails can be estimated nonparametrically along with the (common) extreme value index. Joint asymptotic normality of both estimators is shown; they are asymptotically independent. We develop tests for the proportionality function and for the validity of the model. We show through simulations the good performance of tests for tail homoscedasticity. The results are applied to stock market returns. A main tool is the weak convergence of a weighted sequential tail empirical process.

Key words and phrases. Extreme value statistics, functional limit theorems, scale, sequential tail empirical process, non-identical distributions.

MSC2010 subject classifications. Primary 62G32, 62G05, 62G10, 62G20, 62G30; secondary 60G70, 60F17.

JEL Codes. C12, C13, C14.

1 Introduction

Classical extreme value analysis makes statistical inference on the tail region of a probability distribution, based on independent and identically distributed (i.i.d.) observations. Nevertheless, the observed data may violate the i.i.d. assumption. Two potential deviations may occur: the observations may exhibit serial dependence or they may be drawn from different distributions. In this paper we consider the latter situation and develop extreme value statistics to handle the case when observations are drawn from different distributions. It will turn out that extreme value statistics goes through under mild variation of the underlying distributions and that we can quantify this variation which reflects the frequency of extreme events.

We consider the following model. At “time” points $i = 1, \dots, n$ we have independent observations $X_1^{(n)}, \dots, X_n^{(n)}$ following various continuous distribution functions $F_{n,1}, \dots, F_{n,n}$, that share a common right endpoint $x^* = \sup \{x : F_{n,i}(x) < 1\} \in (-\infty, \infty]$, and there exist a continuous distribution function F with the same right endpoint and a continuous, positive function c defined on $[0, 1]$ such that

$$\lim_{x \rightarrow x^*} \frac{1 - F_{n,i}(x)}{1 - F(x)} = c\left(\frac{i}{n}\right), \quad (1.1)$$

uniformly for all $n \in \mathbb{N}$ and all $1 \leq i \leq n$ (see de Haan et al. (2011)). To make the function c uniquely defined we impose the condition

$$\int_0^1 c(s) ds = 1. \quad (1.2)$$

We call the above situation *heteroscedastic extremes* analogous to the concept of heteroscedasticity and call c the *skedasis function*. It characterizes the trend in extremes. For example $c \equiv 1$ corresponds to no trend or “homoscedastic extremes”. Notice that condition (1.1) assumes proportionality of only the tail parts of the underlying distribution functions. Hence, we do not impose any assumption on the central parts of the distributions. It describes a flexible nonparametric model that allows for different scales in the tails, as explained below.

In addition, we assume, as in classical extreme value analysis, that F belongs to the domain of attraction of a generalized extreme value distribution. That means, there exists a real number γ and a positive scale function a such that, for all $x > 0$,

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \quad (1.3)$$

where $U := \left(\frac{1}{1-F}\right)^{\leftarrow}$ and \leftarrow denotes the left-continuous inverse function. The index γ is the extreme value index. Write also $U_{n,i} := \left(\frac{1}{1-F_{n,i}}\right)^{\leftarrow}$. Combining the domain of attraction condition with (1.1), it can be shown that

$$\lim_{t \rightarrow \infty} \frac{U_{n,i}(tx) - U_{n,i}(t)}{a(t) \left(c\left(\frac{i}{n}\right)\right)^\gamma} = \frac{x^\gamma - 1}{\gamma}. \quad (1.4)$$

Hence, all $F_{n,i}$ belong to the domain of attraction of the same extreme value distribution. They have the same extreme value index γ but (for $\gamma \neq 0$) different scale functions a , as in (1.3), that is, the impact of the variation in the function c is on the scale of extremes instead of on the extreme value index. If $\gamma = 0$ the impact is on the location only.

In the sequel we will restrict ourselves to the heavy-tailed case, i.e., $\gamma > 0$. This is done in view of applications in finance. Then, $x^* = \infty$ and the domain of attraction condition (1.3) simplifies to

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\gamma. \quad (1.5)$$

Then the analogue of (1.4) is

$$\lim_{t \rightarrow \infty} \frac{U_{n,i}(tx)}{U(t) \left(c\left(\frac{i}{n}\right)\right)^\gamma} = x^\gamma.$$

In this paper we make the following contributions. First, we propose a nonparametric estimator on the integrated skedasis function $C(s) := \int_0^s c(u) du$, for $s \in [0, 1]$, and establish its asymptotic behavior. Moreover, we show that the Hill estimator can still be successfully applied to estimate the extreme value index γ , even though the observations are drawn from different distributions. The joint asymptotic distribution of both estimators is established. The estimators of γ and C are asymptotically independent. Second, we test hypotheses on (the presence of) heteroscedastic extremes. The null hypothesis is $c = c_0$ for some given skedasis function c_0 . In particular, rejecting the null hypothesis $c \equiv 1$, confirms that extreme events in a certain period of history are more frequent than in other periods. Third, for application purposes, we provide estimators of c and of high quantiles corresponding to $F_{n,i}$. In applications, the evolution in time of the high quantile estimates quantifies the impact of heteroscedasticity on the magnitude of extreme events. All of this is presented in Section 2. In Section 3, we validate our model by testing if the extreme value index is constant over time. In Section 4 we present a small simulation study and apply our results to financial data.

We sketch how we handle heteroscedastic extremes statistically. Consider $X_i^{(n)}$ for $i = 1, \dots, n$. We impose a high threshold. Then the (local) frequency of the exceedances over the threshold reflects the (local) value of the skedasis function whereas the magnitude of the exceedances reflects the value of the extreme value index.

A crucial tool for developing the asymptotic theory is the sequential tail empirical process (STEP), based on non-identically distributed observations. similar to the sequential empirical process (see, e.g., Section 3.5 in Shorack and Wellner (1986)), the STEP is a bivariate process with one coordinate denoting time and the other one magnitude. We prove, in Section 5, the weighted convergence of the STEP to a bivariate Wiener process. Since all our estimators and test statistics can be written as functionals of the STEP, their statistical properties follow from this result. The asymptotic theory for the STEP is of independent interest and can be used for analyzing other statistical procedures for heteroscedastic extremes. In particular, it can be used for proving asymptotic theory for other extreme value index estimators (even when γ is not positive). Also, other tests for testing on heteroscedastic extremes or constant extreme value index can be analyzed using the STEP. Our test statistics for constant extreme value index are only the more straightforward candidates.

Our study is comparable with other deviations from the i.i.d. assumption in extreme value analysis. In the direction of allowing serial dependence, Leadbetter et al. (1983) develops the probability theory on extremes of stationary weakly dependent time series. Hsing (1991), Drees (2000) and recently Drees and Rootzén (2010), further develop statistical tools to handle extreme events for weakly dependent observations. In all these studies, the observations are assumed to be identically distributed. In the direction of allowing heteroscedastic extremes, the early contribution Meijzler (1956) provides a probabilistic theory based on independent, non-identically distributed random variables. As to statistical analysis of heteroscedastic extremes, a few studies have explored a trend in the parameters of some *limit* distributions in EVT. Davison and Smith (1990) consider a linear trend in both shape and scale parameters of generalized Pareto distributions (GPD), while Coles (2001) deals with a log-linear trend in the scale parameter of GPDs. No asymptotic analysis of the estimators was provided in these studies. Two other studies have provided estimators on trends in extremes with asymptotic properties. Hall and Tajvidi (2000) estimate nonparametric trends in parameters of GPDs and general-

ized extreme value distributions by locally parametrizing the trend. They establish the asymptotic behavior of the estimators under locally constant or locally linear trends. Differently, de Haan et al. (2011) considers a similar model as in our study, but concentrates on specific parametric trends and requires a large number of observations at any time point. Compared to all existing studies on heteroscedastic extremes, our approach differs in one or more of the following three aspects: we deal with an extreme value analysis based on the domain of attraction rather than the limit situation; we do not impose any (local) parametric model on the skedasis function; we provide asymptotic properties of the estimators.

This paper also contributes to the literature on testing whether the extreme value index is constant over time. For example, Quintos et al. (2001) investigates whether the extreme value index of financial data is time invariant. The test statistics therein focus only on tail behavior of observations. The asymptotic theory of the tests statistics assumes that the observations are i.i.d., which is much more strict than the targeted null hypothesis that the extreme value index is invariant over time. Consequently, the testing procedure there would reject in case of heteroscedastic extremes where in fact the extreme value index is constant. In contrast, our test considers the much larger heteroscedastic extremes model as the null hypothesis.

2 Estimation and testing within the heteroscedastic extremes model

In this section, we consider statistical inference on the skedasis function c and also estimation of the common extreme value index γ . We begin with estimating the integrated function c , defined by $C(s) := \int_0^s c(u)du$, for $s \in [0, 1]$. Intuitively, by focusing on the observations above a high threshold, the function C should be proportional to the number of exceedances of the threshold in the first $[ns]$ observations. This leads to the following estimator. Order the observations $X_1^{(n)}, \dots, X_n^{(n)}$ as $X_{n,1} \leq \dots \leq X_{n,n}$. For a suitable intermediate sequence $k = k(n)$, that is,

$$\lim_{n \rightarrow \infty} k = \infty, \quad \lim_{n \rightarrow \infty} \frac{k}{n} = 0, \quad (2.1)$$

we define the estimator

$$\hat{C}(s) := \frac{1}{k} \sum_{i=1}^{[ns]} 1_{\{X_i^{(n)} > X_{n,n-k}\}}. \quad (2.2)$$

When the observations are all different, the estimator can be written in terms of the ranks $R_{n,i} = \sum_{j=1}^n 1_{\{X_i^{(n)} \geq X_j^{(n)}\}}$, $1 \leq i \leq n$, as follows,

$$\hat{C}(s) = \frac{1}{k} \sum_{i=1}^{[ns]} 1_{\{R_{n,i} > n-k\}}.$$

Next we define the Hill estimator as usual

$$\hat{\gamma}_H := \frac{1}{k} \sum_{j=1}^k \log X_{n,n-j+1} - \log X_{n,n-k}, \quad (2.3)$$

but note that is not yet clear that this is a proper estimator of γ in case of heteroscedastic extremes.

In order to prove the asymptotic normality of these estimators, we need second-order conditions quantifying the speed of convergence in (1.1) and (1.5) respectively. Firstly, suppose that there exists a positive, eventually decreasing function A_1 , with $\lim_{t \rightarrow \infty} A_1(t) = 0$, such that, as $x \rightarrow \infty$,

$$\sup_{n \in \mathbb{N}} \max_{1 \leq i \leq n} \left| \frac{1 - F_{n,i}(x)}{1 - F(x)} - c\left(\frac{i}{n}\right) \right| = O\left(A_1\left(\frac{1}{1 - F(x)}\right)\right). \quad (2.4)$$

Secondly, suppose that there exists a function A_2 and a $\rho < 0$ such that, as $t \rightarrow \infty$, $A_2(t)$ has either positive or negative sign, $A_2(t) \rightarrow 0$, and for any $x > 0$,

$$\lim_{t \rightarrow \infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A_2(t)} = x^\gamma \frac{x^\rho - 1}{\rho}, \quad (2.5)$$

see de Haan and Stadtmüller (1996). We require, as $n \rightarrow \infty$,

$$\sqrt{k}A_1(n/k) \rightarrow 0 \quad \text{and} \quad \sqrt{k}A_2(n/k) \rightarrow 0. \quad (2.6)$$

We further assume that

$$\lim_{n \rightarrow \infty} \sqrt{k} \sup_{|u-v| \leq 1/n} |c(u) - c(v)| = 0. \quad (2.7)$$

Assumption (2.7) is rather weak: if c is Lipschitz continuous of order at least $1/2$, it is a direct consequence of the fact that $k/n \rightarrow 0$, as $n \rightarrow \infty$.

The following theorem on the joint asymptotic normality of \hat{C} and $\hat{\gamma}_H$ is our main result.

Theorem 2.1 *Suppose conditions (1.2), (2.1), (2.4), (2.5), (2.6), and (2.7), hold. Then, under a Skorokhod construction, we have that, as $n \rightarrow \infty$,*

$$\sup_{0 \leq s \leq 1} \left| \sqrt{k}(\hat{C}(s) - C(s)) - B(C(s)) \right| \rightarrow 0 \text{ a.s.}$$

and

$$\sqrt{k}(\hat{\gamma}_H - \gamma) \rightarrow \gamma N_0 \text{ a.s.,}$$

with B a standard Brownian bridge and N_0 a standard normal random variable. In addition, B and N_0 are independent.

Remark Observe that the asymptotic variance of the Hill estimator $\hat{\gamma}_H$ does not depend on c , hence it is the same as in the i.i.d. case ($c \equiv 1$). Recall that $\hat{\gamma}_H$ is based on the order statistics and \hat{C} on the ranks. In the i.i.d. case the vector of order statistics and the vector of ranks are independent, yielding the independence of $\hat{\gamma}_H$ and \hat{C} . In the case of heteroscedastic extremes we do not have the independence of ranks and order statistics, nevertheless we have asymptotic independence of $\hat{\gamma}_H$ and \hat{C} . From the proofs (Sections 5 and 6) it follows that the asymptotic independence of $\hat{\gamma}$ and \hat{C} also holds for the other estimators in use for γ (even for the broader case $\gamma \in \mathbb{R}$), that is, the estimator of the trend in extremes and that of the extreme value index are asymptotically independent. In fact, the asymptotic theory for \hat{C} does not require the extreme value condition (1.3).

Next, we present an estimator of the function c by using a kernel estimation method. Let G be a continuous kernel function on $[-1, 1]$ such that $\int_{-1}^1 G(s)ds = 1$; set $G(s) = 0$ for $|s| > 1$. Let $h := h_n > 0$ be a bandwidth such that $h \rightarrow 0$ and $kh \rightarrow \infty$, as $n \rightarrow \infty$. Then, the function c can be estimated nonparametrically by

$$\hat{c}(s) := \frac{1}{kh} \sum_{i=1}^n 1_{\{X_i > X_{n,n-k}\}} G\left(\frac{s - \frac{i}{n}}{h}\right).$$

This estimator is similar to the usual kernel density estimator. An example of a kernel function G is the biweight kernel $15(1 - x^2)^2/16$ on $[-1, 1]$. This kernel will be used in the application in Section 4.

Instead of estimating c , we can also test the null hypothesis that $c = c_0$ for some given function c_0 . This is equivalent to testing $C = C_0$ with $C_0(s) := \int_0^s c_0(u)du$. An important example is testing the null hypothesis $c \equiv 1$, which corresponds to testing C is the identity function on $[0, 1]$. By rejecting this null hypothesis, we can conclude that

extreme events in a certain period of history are more frequent than in other periods. We consider a Kolmogorov-Smirnov type test statistic

$$T_1 := \sup_{0 \leq s \leq 1} \left| \hat{C}(s) - C_0(s) \right|$$

and a Cramér-von Mises type test statistic

$$T_2 := \int_0^1 \left(\hat{C}(s) - C_0(s) \right)^2 dC_0(s).$$

The following direct corollary to Theorem 2.1 gives the asymptotic distributions of these two test statistics under H_0 .

Corollary 2.2 *Assume that the conditions of Theorem 2.1 hold with $c = c_0$. Then, as $n \rightarrow \infty$,*

$$\begin{aligned} \sqrt{k}T_1 &\xrightarrow{d} \sup_{0 \leq s \leq 1} |B(s)|, \\ kT_2 &\xrightarrow{d} \int_0^1 B^2(s)ds, \end{aligned}$$

with B a standard Brownian bridge.

Observe that the limiting random variables have well-known probability distributions that do not depend on c_0 . Also, the domain of attraction condition on F does not play a role and thus these tests can be applied to a broader class of probability distributions.

Finally, we present how to estimate high quantiles at a time point i when having heteroscedastic extremes. High quantiles are the quantiles $U_{n,i}(1/p)$ with very small tail probability p , where p can be even less than $1/n$. According to (1.1), we have

$$p = 1 - F_{n,i}(U_{n,i}(1/p)) \approx c \left(\frac{i}{n} \right) (1 - F(U_{n,i}(1/p))).$$

Hence we obtain

$$U_{n,i} \left(\frac{1}{p} \right) \approx U \left(\frac{c(i/n)}{p} \right).$$

Therefore, to estimate $U_{n,i}(1/p)$ we combine the estimator of the skedasis function c with the quantile estimator corresponding to the distribution function F (cf. Weissman (1978)) and obtain

$$\widehat{U_{n,i}(1/p)} = X_{n,n-k} \left(\frac{k\hat{c}(i/n)}{np} \right)^{\hat{\gamma}_H}.$$

The high quantile estimator can be extended to *forecasting* tail risks, that is, we intend to estimate the high quantile of an unobserved random variable in the next period, $X_{n+1}^{(n)}$. Extending the function c continuously in a right neighborhood of 1 and incorporating time point $i = n + 1$ in (1.1), leads to the following estimator of the high quantile $U_{n,n+1}(1/p)$ of the unobserved $X_{n+1}^{(n)}$:

$$\widehat{U_{n,n+1}(1/p)} = X_{n,n-k} \left(\frac{k\hat{c}(1)}{np} \right)^{\hat{\gamma}_H}.$$

Since the estimator involves \hat{c} at the boundary point 1, we recommend using boundary kernels, see e.g. Jones (1993).

3 Testing if the extreme value index is constant

Here we consider the validity of our model. In particular we test if the extreme value index γ is constant over time. The idea is to estimate the γ from subsamples and compare the estimates. Concretely, we write $\hat{\gamma}_{(s_1, s_2]}$ for the Hill estimator based on $X_{[ns_1]+1}, \dots, X_{[ns_2]}$, for any $0 \leq s_1 < s_2 \leq 1$. Recall that when estimating γ from the full sample, we use the highest $k + 1$ observations. Correspondingly, the number of high observations used in $\hat{\gamma}_{(s_1, s_2]}$ has to reflect the heteroscedasticity in extremes. We would like to choose $k_{(s_1, s_2]}^* := k(C(s_2) - C(s_1))$, which is proportional to the frequency of having exceedances in this subsample. In practice, we estimate it with $k_{(s_1, s_2]} := k(\hat{C}(s_2) - \hat{C}(s_1))$. The following theorem shows the joint asymptotic behavior of these partial Hill estimators. The proof is deferred to Section 6.

Theorem 3.1 *Assume that the conditions of Theorem 2.1 hold. Then, under a Skorokhod construction, we have that for any $\delta > 0$, as $n \rightarrow \infty$,*

$$\sup_{0 \leq s_1 < s_2 \leq 1, s_2 - s_1 \geq \delta} \left| \sqrt{k}(\hat{\gamma}_{(s_1, s_2]} - \gamma) - \gamma \frac{W(C(s_2)) - W(C(s_1))}{C(s_2) - C(s_1)} \right| \rightarrow 0 \text{ a.s.},$$

where W is a standard Wiener process on $[0, 1]$. In addition, the process W and the Brownian bridge B from Theorem 2.1 are independent and $W(1)$ is equal to N_0 there.

The first test compares all partial Hill estimators such that $\hat{C}(s_2) - \hat{C}(s_1) \geq \delta$, for some given $\delta > 0$, to the one using the full sample, $\hat{\gamma}_{(0,1]} = \hat{\gamma}_H$. The test statistic is

$$T_3 := \sup_{0 \leq s_1 < s_2 \leq 1, \hat{C}(s_2) - \hat{C}(s_1) \geq \delta} \left| \frac{\hat{\gamma}_{(s_1, s_2]}}{\hat{\gamma}_H} - 1 \right|.$$

Alternatively, we consider a test statistic with a limited number of partial Hill estimators. Divide the sample into m blocks, with $m > 1$ fixed. The cutoff points of the blocks are $l_1 \leq l_2 \leq \dots \leq l_{m-1}$ with $l_j := \sup \left\{ s : \hat{C}(s) \leq j/m \right\}$; set $l_0 = 0$ and $l_m = 1$. We use the partial Hill estimator $\hat{\gamma}_{(l_{j-1}, l_j]}$ as above, but use the highest $[k/m] + 1$ observations in each subsample, since, by construction, $\hat{C}(l_j) - \hat{C}(l_{j-1})$ is approximately $1/m$ for each j . Now define the test statistic as

$$T_4 := \frac{1}{m} \sum_{j=1}^m \left(\frac{\hat{\gamma}_{(l_{j-1}, l_j]}}{\hat{\gamma}_H} - 1 \right)^2.$$

Corollary 3.2 *Assume that the conditions of Theorem 2.1 hold. Then, we have that, as $n \rightarrow \infty$,*

$$\begin{aligned} \sqrt{k}T_3 &\xrightarrow{d} \sup_{0 \leq s_1 < s_2 \leq 1, s_2 - s_1 \geq \delta} \left| \frac{W(s_2) - W(s_1)}{s_2 - s_1} - W(1) \right|, \\ kT_4 &\xrightarrow{d} \chi_{m-1}^2, \end{aligned}$$

with W a standard Wiener process.

The proof is deferred to Section 6. Observe that the limits do not depend on c or γ .

4 Simulations and application

In this section we will first examine, through simulations, the finite sample behavior of the two tests on the skedasis function of Section 2 (Subsection 4.1). Next, in Subsection 4.2, we will apply all the tests to check whether the extreme value index (T_3, T_4) and the skedasis function (T_1, T_2) of a stock market return series are invariant over time and we will also estimate the skedasis function.

4.1 Simulations

We consider four data generating processes (DGPs) as follows.

DGP 1: Observations are i.i.d. and follow the standard Fréchet distribution, i.e. $F_{i,n}^{(1)}(x) = \exp(-1/x)$, for $x > 0$. Here $c \equiv 1$.

DGP 2: Observations are independent, with observation i following a rescaled Fréchet distribution: $F_{i,n}^{(2)}(x) = \exp\left(-\frac{0.5+i/n}{x}\right)$, for $x > 0$. Here $c(s) = 0.5 + s$, for $s \in [0, 1]$.

DGP 3: Observations are independent, with observation i following a rescaled Fréchet distribution: $F_{i,n}^{(3)}(x) = \exp\left(-\frac{c(i/n)}{x}\right)$, for $x > 0$, with $c(s) = 2s + 0.5$, for $s \in [0, 0.5]$, $c(s) = -2s + 2.5$ for $s \in (0.5, 1]$,

DGP 4: Observations are independent, with observation i following a rescaled Fréchet distribution: $F_{i,n}^{(3)}(x) = \exp\left(-\frac{c(i/n)}{x}\right)$, for $x > 0$, with $c(s) = 0.8$, for $s \in [0, 0.4] \cup [0.6, 1]$, $c(s) = 20s - 7.2$ for $s \in (0.4, 0.5]$, $c(s) = -20s + 12.8$ for $s \in (0.5, 0.6)$.

For each DGP, we simulate 1000 samples of size $n = 5000$. We apply the two tests of Section 2 to test whether there exist heteroscedastic extremes ($H_0 : c \equiv 1$), with $k = 400$. For each significance level (1%, 5% and 10%), we show in Table 1 the total number (out of 1000) of rejections for each DGP. We see that both tests perform well, both under the

α	1%		5%		10%	
Test	T_1	T_2	T_1	T_2	T_1	T_2
DGP 1	8	12	44	47	95	98
DGP 2	990	998	998	999	1000	1000
DGP 3	455	570	838	921	941	987
DGP 4	663	521	930	903	979	978

Table 1: Number of rejections out of 1000 simulated datasets.

null hypothesis (DGP1) and under the alternative (DGPs 2-4). In particular the power is high in most cases. Test 2 performs somewhat better for global deviations from the null hypothesis, whereas Test 1 detects the spike alternative a bit better.

4.2 Application

We apply the proposed estimators and testing procedures to address the question “Are financial crises nowadays more frequent than before?”. For that purpose, we collect daily loss returns of the S&P 500 index from 1988 to 2012 as an indicator for the status of the US financial market over this period. It has been documented in the empirical finance literature that the downside of equity returns follows heavy-tailed distributions, see e.g. Jansen and de Vries (1991) and Kearns and Pagan (1997). Assuming that the

loss returns on each day follow, possibly different, heavy-tailed distributions as in (1.1) and (1.5), we test whether the extreme value index of the loss returns is invariant over time. If not rejected, we further test whether the skedasis function is invariant over time.

We start with analyzing the full sample from 1988 to 2012, consisting of 6302 observations (2926 days with losses) and use $k = 160$. Tests 3 (with $\delta = 1/4$) and 4 (with $m = 4$) both yield p -values that are virtually zero. Hence, we strongly reject the null that the extreme value index is invariant over time. We do not need to further investigate the skedasis function as our model is not valid for this dataset.

The observed structural change in the extreme value index might be attributed to the recent financial crisis. Therefore we continue with a 20-years subsample from 1988 to 2007, consisting of 5043 observations (2348 days with losses). This excludes the recent financial crisis (and the so-called “black Monday” in 1987), but nevertheless includes other crisis events such as the burst of the internet bubble at the beginning of the 21st century. We test again the null that the extreme value index is invariant during this period using $k = 130$. Tests 3 and 4 yield p -values 0.98 and 0.76, respectively. Hence, we do not reject the null of constant extreme value index. In other words, the crisis magnitude, measured by the extreme value index, is not varying during this period.

We further test whether the skedasis function is constant in the subsample from 1988 to 2007. Both Tests 1 and 2 report strong evidence rejecting the null (p -values are virtually zero). Hence, although the magnitude remains at a constant level, the tail frequency is time varying during this period. We apply our kernel estimator \hat{c} of Section 2, with the biweight kernel $15(1 - x^2)^2/16$ and bandwidth $h = 0.1$, to estimate the function c . The estimate \hat{c} is plotted in the upper panel of Figure 1. We observe the peak of the skedasis function in the period from 2001 to 2002, which reflects the burst of the internet bubble. We conclude that the tail risk during these two years is higher than that during other periods. Note that at the end of the period, the skedasis function c increases again, even though we use only data up to the end of 2007, before the financial crisis erupts.

We check the robustness of our results using *weekly* equity returns. The daily equity return series may suffer from serial dependence such as volatility clustering, which violates our assumption on independence. Such serial dependence is at least much weaker in weekly returns. We repeat our analysis for the weekly loss returns in the subsample from 1988 to 2007, consisting of 1043 observations (454 weeks with losses). Using $k = 60$,

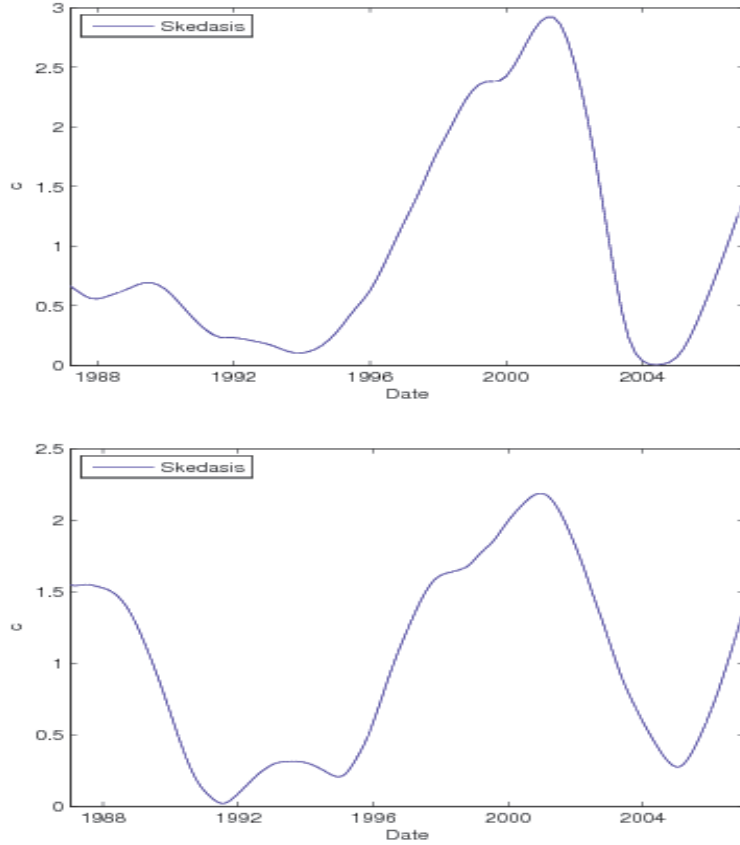


Figure 1: The estimated skedasis function c based on daily (upper panel) and weekly (lower panel) loss returns of the S&P 500 index from 1988 till 2007.

Tests 3 and 4 yield p -values 0.21 and 0.18, respectively. Hence, we do not reject the null of constant extreme value index. In addition, Tests 1 and 2 yield p -values 0.01 and 0.03, respectively, which provides evidence that the tail frequency is time varying during this period. Lastly, with the same kernel estimator \hat{c} , we estimate the skedasis function c during this period (lower panel of Figure 1). We see from both the quantitative and qualitative analysis that our results are robust when changing the frequency of the data.

5 The STEP

The proofs of the theorems in Sections 2 and 3 are based on a specific tool: the sequential tail empirical process (STEP). In this section, we define the STEP and study its asymptotic properties. Recall that the function c is positive and continuous on $[0, 1]$. Thus, there exist positive numbers b and d such that $0 < b < c(s) < d$, for all $s \in [0, 1]$.

Define the sequential empirical distribution function as

$$F_n(x, s) := \frac{1}{n} \sum_{i=1}^{[ns]} 1_{\{X_i^{(n)} \leq x\}}, \quad x < x^*.$$

Since we are interested in the right tail of the distribution, we further define the sequential empirical survival function as

$$\bar{F}_n(x, s) := \frac{1}{n} \sum_{i=1}^{[ns]} 1_{\{X_i^{(n)} > x\}} = \frac{[ns]}{n} - F_n(x, s), \quad x < x^*.$$

Next, we deal with the tail region corresponding to $x = U\left(\frac{n}{kt}\right)$, for $0 \leq t \leq 1$, where k satisfies (2.1). We approximate the mean and variance of $\bar{F}_n\left(U\left(\frac{n}{kt}\right), s\right)$ as follows. From the limit relation (1.1),

$$\begin{aligned} \mathbb{E} \bar{F}_n\left(U\left(\frac{n}{kt}\right), s\right) &= \frac{1}{n} \sum_{i=1}^{[ns]} \left(1 - F_{n,i}\left(U\left(\frac{n}{kt}\right)\right)\right) \\ &\approx \frac{1}{n} \sum_{i=1}^{[ns]} c\left(\frac{i}{n}\right) \left(1 - F\left(U\left(\frac{n}{kt}\right)\right)\right) \\ &\approx \frac{kt}{n} C(s). \end{aligned}$$

Similarly, as $n \rightarrow \infty$, we get the approximation of the variance as

$$\text{Var}\left(\bar{F}_n\left(U\left(\frac{n}{kt}\right), s\right)\right) = \frac{1}{n^2} \sum_{i=1}^{[ns]} \left(1 - F_{n,i}\left(U\left(\frac{n}{kt}\right)\right)\right) F_{n,i}\left(U\left(\frac{n}{kt}\right)\right) \approx \frac{kt}{n^2} C(s) = O\left(\frac{k}{n^2}\right).$$

Normalizing $\bar{F}_n\left(U\left(\frac{n}{kt}\right), s\right)$ with the approximations of its expectation and variance, we define the *sequential tail empirical process* (STEP) as

$$\begin{aligned} \mathbb{F}_n(t, s) &:= \sqrt{\frac{n^2}{k}} \left(\bar{F}_n\left(U\left(\frac{n}{kt}\right), s\right) - \frac{kt}{n} C(s) \right) \\ &= \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^{[ns]} 1_{\{X_i^{(n)} > U\left(\frac{n}{kt}\right)\}} - t C(s) \right). \end{aligned}$$

We shall prove that under proper conditions, the STEP converges to a Wiener process in a proper function space.

We start with considering the “simple” case where F is a standard uniform distribution function and the limit relation in (1.1) is exact. That is, for all $1 \leq i \leq n$, $1 - F_{n,i}(x) = c\left(\frac{i}{n}\right)(1 - x)$, for $x \in [1 - \frac{1}{c(\frac{i}{n})}, 1]$. In that case, each $X_i^{(n)}$ follows a uniform distribution

on $[1 - \frac{1}{c(\frac{i}{n})}, 1]$. Hence, we can write $X_i^{(n)} = 1 - \frac{U_i}{c(\frac{i}{n})}$, where the U_i are i.i.d. uniform-[0,1] random variables. The STEP in this special case is then written as

$$\mathbb{S}_n(t, s) = \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^{[ns]} 1_{\{U_i < c(\frac{i}{n}) \frac{kt}{n}\}} - tC(s) \right).$$

We call it the *simple STEP*.

We first establish the asymptotic behavior of the simple STEP. Firstly, we extend the definition of the simple STEP to $(t, s) \in D := (0, 2] \times [0, 1]$ with the same formula. Secondly, we define a weight function $q(t) = t^\eta$ for any $0 \leq \eta < 1/2$. Then, we have the following proposition.

Proposition 5.1 *Suppose k satisfies (2.1) and (2.7). Under a Skorokhod construction, there exists a standard bivariate Wiener process \tilde{W} on D , that is, \tilde{W} is a mean zero Gaussian process with*

$$\text{Cov}(\tilde{W}(t_1, s_1), \tilde{W}(t_2, s_2)) = (t_1 \wedge t_2)(s_1 \wedge s_2), \text{ for } (t_1, s_1), (t_2, s_2) \in D,$$

such that, as $n \rightarrow \infty$,

$$\sup_{(t,s) \in D} \frac{1}{q(t)} \left| \mathbb{S}_n(t, s) - \tilde{W}(t, C(s)) \right| \rightarrow 0 \text{ a.s.}$$

The proof of this proposition requires the following two lemmas.

Lemma 5.2 *For independent, uniform-[0,1] random variables V_1, \dots, V_n , define*

$$K_n(t, s) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[ns]} [1_{\{V_i < t\}} - t], \quad 0 \leq t, s \leq 1.$$

Let K denote a Kiefer process on $[0, 1]^2$, that is, K is a mean zero Gaussian process with

$$\text{Cov}(K(t_1, s_1), K(t_2, s_2)) = (t_1 \wedge t_2 - t_1 t_2)(s_1 \wedge s_2), \text{ for } (t_1, s_1), (t_2, s_2) \in [0, 1]^2$$

Then, we have, under a Skorokhod construction, as $n \rightarrow \infty$,

$$\sup_{0 < t \leq 1, 0 \leq s \leq 1} \frac{1}{q(t)} |K_n(t, s) - K(t, s)| \rightarrow 0 \text{ a.s.}$$

Lemma 5.3 *Suppose Z_1, \dots, Z_n are independent random variables with Bernoulli distributions: $P(Z_i = 1) = 2c(\frac{i}{n}) \frac{k}{n}$, with k satisfying (2.1) and (2.7). Define the partial sum process as*

$$N_n(s) = \sum_{i=1}^{[ns]} Z_i.$$

Then, under a Skorokhod construction, there exists a standard Wiener process W_0 on $[0, 2]$, such that, as $n \rightarrow \infty$,

$$\sup_{0 \leq s \leq 1} \left| \sqrt{k} \left(\frac{N_n(s)}{k} - 2C(s) \right) - W_0(2C(s)) \right| \rightarrow 0 \text{ a.s.}$$

The first lemma follows from Theorem 2.12.1 in van der Vaart and Wellner (1996) in combination with the Chibisov-O'Reilly theorem (see p. 462 in Shorack and Wellner (1986)). In fact, the lemma holds with any non-decreasing continuous function $q : [0, 2] \rightarrow (0, \infty)$ such that

$$\int_0^2 u^{-1} \exp(-\lambda q^2(u)/u) du < \infty,$$

for all $\lambda > 0$.

Proof of Lemma 5.3 We apply Theorem 2.12.6 in van der Vaart and Wellner (1996) with $Y_{ni} = \frac{1}{\sqrt{k}}(Z_i - E Z_i)$, Q_{ni} being equal to the Dirac measure at i/n and Q being equal to a measure on $[0, 1]$ such that $Q([0, s]) = 2C(s)$. We have that, under a Skorokhod construction, there exists a standard Wiener process W_0 on $[0, 2]$, such that, as $n \rightarrow \infty$,

$$\sup_{0 \leq s \leq 1} \left| \sqrt{k} \left(\frac{N_n(s)}{k} - 2 \frac{1}{n} \sum_{i=1}^{[ns]} c\left(\frac{i}{n}\right) \right) - W_0(2C(s)) \right| \rightarrow 0 \text{ a.s.}$$

The lemma is proved provided that $\sup_{0 \leq s \leq 1} \sqrt{k} \left| \frac{1}{n} \sum_{i=1}^{[ns]} c\left(\frac{i}{n}\right) - C(s) \right| \rightarrow 0$ as $n \rightarrow \infty$, which follows from (2.7). \square

Proof of Proposition 5.1 First, we construct n independent uniform-[0,1] random variables U_1, U_2, \dots, U_n in a special way. Recall that d is the upper bound of the function c . For n such that $\frac{n}{k} > 2d$, let Z_i , $1 \leq i \leq n$ be independent random variables following Bernoulli distributions with $P(Z_i = 1) = 2c\left(\frac{i}{n}\right) \frac{k}{n}$. Let V_j , $1 \leq j \leq n$, be independent uniform-[0,1] random variables, independent of the Z_i . We combine these $2n$ random variables to construct the U_i . Each Z_i is matched with a V_j , where the random index j is defined as follows (recall the notation of Lemma 5.3):

$$j = \begin{cases} N_n\left(\frac{i-1}{n}\right) + 1 & \text{if } Z_i = 1, \\ N_n(1) + i - N_n\left(\frac{i-1}{n}\right) & \text{if } Z_i = 0. \end{cases}$$

That is, we assign the first $N_n(1)$ random variables V_j to the Z_i with $Z_i = 1$, and then assign the rest of the V_j to the Z_i with $Z_i = 0$. Then we construct

$$U_i = Z_i 2c\left(\frac{i}{n}\right) \frac{k}{n} V_j + (1 - Z_i) \left\{ 2c\left(\frac{i}{n}\right) \frac{k}{n} + \left(1 - 2c\left(\frac{i}{n}\right) \frac{k}{n}\right) V_j \right\}, i = 1, \dots, n.$$

It is straightforward to verify that U_1, \dots, U_n are independent uniform-[0,1] random variables.

We base our simple STEP on these U_i . We then get (recalling the notation of Lemma 5.2):

$$\begin{aligned}
\mathbb{S}_n(t, s) &= \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^{\lfloor ns \rfloor} 1_{\{U_i < c(\frac{i}{n}, \frac{kt}{n})\}} - tC(s) \right) \\
&= \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^{N_n(s)} 1_{\{V_i < \frac{t}{2}\}} - tC(s) \right) \\
&= \left(\frac{N_n(1)}{k} \right)^{1/2} \frac{1}{\sqrt{N_n(1)}} \sum_{i=1}^{N_n(s)} \left(1_{\{V_i < \frac{t}{2}\}} - \frac{t}{2} \right) + \frac{t}{2} \sqrt{k} \left(\frac{N_n(s)}{k} - 2C(s) \right) \\
&= \left(\frac{N_n(1)}{k} \right)^{1/2} K_{N_n(1)} \left(\frac{t}{2}, \frac{N_n(s)}{N_n(1)} \right) + \frac{t}{2} \sqrt{k} \left(\frac{N_n(s)}{k} - 2C(s) \right) \\
&=: I_1(t, s) + I_2(t, s).
\end{aligned} \tag{5.1}$$

Observe that the two sequences of processes $\{K_m\}$ and $\{N_n\}$ are independent, and hence their limits K and W_0 are independent. We have

$$\begin{aligned}
&\frac{1}{q(t)} \left| I_1(t, s) - \sqrt{2}K \left(\frac{t}{2}, C(s) \right) \right| \\
&\leq \left(\frac{N_n(1)}{k} \right)^{1/2} \frac{1}{q(t)} \left| K_{N_n(1)} \left(\frac{t}{2}, \frac{N_n(s)}{N_n(1)} \right) - K \left(\frac{t}{2}, C(s) \right) \right| \\
&\quad + \frac{|K(\frac{t}{2}, C(s))|}{q(t)} \left| \left(\frac{N_n(1)}{k} \right)^{1/2} - \sqrt{2} \right|.
\end{aligned}$$

Now it readily follows from Lemmas 5.2 and 5.3 that

$$\sup_{(t,s) \in D} \frac{1}{q(t)} \left| I_1(t, s) - \sqrt{2}K \left(\frac{t}{2}, C(s) \right) \right| \rightarrow 0 \text{ a.s.} \tag{5.2}$$

It is immediate from Lemma 5.3 that, as $n \rightarrow \infty$,

$$\sup_{(t,s) \in D} \frac{1}{q(t)} \left| I_2(t, s) - \frac{t}{2} W_0(2C(s)) \right| \rightarrow 0 \text{ a.s.} \tag{5.3}$$

Combining (5.2) and (5.3), yields, as $n \rightarrow \infty$,

$$\sup_{(t,s) \in D} \frac{1}{q(t)} \left| \mathbb{S}_n(t, s) - \left[\sqrt{2}K \left(\frac{t}{2}, C(s) \right) + \frac{t}{2} W_0(2C(s)) \right] \right| \rightarrow 0 \text{ a.s.}$$

Finally write

$$\tilde{W}(t, s) = \sqrt{2}K \left(\frac{t}{2}, s \right) + \frac{t}{2} W_0(2s),$$

and note that \tilde{W} is a standard bivariate Wiener process on D . \square

The following theorem gives the asymptotic behavior of the STEP in the general case, that is, in the setup of Sections 1 and 2.

Theorem 5.4 *Suppose conditions (1.2), (2.1), (2.4), the first part of (2.6), and (2.7), hold. Then, under a Skorokhod construction, there exists a standard bivariate Wiener process \tilde{W} on $[0, 1]^2$ such that, as $n \rightarrow \infty$,*

$$\sup_{0 < t \leq 1, 0 \leq s \leq 1} \frac{1}{q(t)} \left| \mathbb{F}_n(t, s) - \tilde{W}(t, C(s)) \right| \rightarrow 0 \text{ a.s.} \quad (5.4)$$

Proof Denote $U_i = 1 - F_{n,i}(X_i^{(n)})$. Then U_1, \dots, U_n are independent, uniform-[0,1] random variables. We have, almost surely,

$$\mathbb{F}_n(t, s) = \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^{[ns]} 1_{\{U_i < 1 - F_{n,i}(U(\frac{n}{kt}))\}} - tC(s) \right).$$

Condition (2.4) implies that there exists real numbers $x_0 < x^*$ and $\tau > 0$ such that for all $x > x_0$, $n \in \mathbb{N}$ and $1 \leq i \leq n$,

$$c \left(\frac{i}{n} \right) \left(1 - \frac{\tau}{b} A_1 \left(\frac{1}{1 - F(x)} \right) \right) < \frac{1 - F_{n,i}(x)}{1 - F(x)} < c \left(\frac{i}{n} \right) \left(1 + \frac{\tau}{b} A_1 \left(\frac{1}{1 - F(x)} \right) \right).$$

Hence,

$$\mathbb{F}_n^-(t, s) \leq \mathbb{F}_n(t, s) \leq \mathbb{F}_n^+(t, s), \quad (5.5)$$

where

$$\mathbb{F}_n^\pm(t, s) := \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^{[ns]} 1_{\{U_i < c(\frac{i}{n}) \frac{kt}{n} (1 \pm \delta_n)\}} - tC(s) \right),$$

and $\delta_n = \sup_{0 < t \leq 1} \frac{\tau}{b} A_1 \left(\frac{n}{kt} \right) = \frac{\tau}{b} A_1 \left(\frac{n}{k} \right)$.

Next, we study the asymptotic properties of \mathbb{F}_n^+ and \mathbb{F}_n^- . With the standard bivariate Wiener process \tilde{W} of Proposition 5.1, we have

$$\begin{aligned} & \sup_{0 < t \leq 1, 0 \leq s \leq 1} \frac{1}{q(t)} \left| \mathbb{F}_n^+(t, s) - \tilde{W}(t, C(s)) \right| \\ & \leq \sup_{0 < t \leq 1, 0 \leq s \leq 1} \frac{1}{q(t)} \left| \mathbb{S}_n^+(t(1 + \delta_n), s) - \tilde{W}(t(1 + \delta_n), C(s)) \right| \\ & \quad + \sup_{0 < t \leq 1, 0 \leq s \leq 1} \left| \frac{\tilde{W}(t(1 + \delta_n), C(s))}{q(t)} - \frac{\tilde{W}(t, C(s))}{q(t)} \right| \\ & \quad + \sqrt{k} \delta_n \sup_{0 < t \leq 1, 0 \leq s \leq 1} \frac{t}{q(t)} C(s) \\ & =: I_1 + I_2 + I_3. \end{aligned}$$

From Proposition 5.1 it follows that $I_1 \rightarrow 0$ almost surely, as $n \rightarrow \infty$. From the (uniform) continuity of the process $\frac{\tilde{W}(t, C(s))}{q(t)}$, extended to $[0, 2] \times [0, 1]$, we obtain $I_2 \rightarrow 0$, as $n \rightarrow \infty$. Using $\sqrt{k}A_1(n/k) \rightarrow 0$ as $n \rightarrow \infty$, we obtain $I_3 \rightarrow 0$.

Similarly we can show that

$$\sup_{0 < t \leq 1, 0 \leq s \leq 1} \frac{1}{q(t)} \left| \mathbb{F}_n^-(t, s) - \tilde{W}(t, C(s)) \right| \rightarrow 0 \text{ a.s.}$$

Now (5.5) yields (5.4). \square

For Theorem 5.4, we did not use the assumption that F belongs to the domain of attraction. With that assumption, we obtain the following corollary.

Corollary 5.5 *Assume that the conditions in Theorem 2.1 hold. Then, for any $0 \leq \eta < 1/2$ and $x_0 > 0$, under a Skorokhod construction, there exists a standard bivariate Wiener process \tilde{W} on $[0, x_0^{-1/\gamma}] \times [0, 1]$, such that, as $n \rightarrow \infty$,*

$$\sup_{0 \leq s \leq 1, x \geq x_0} x^{\eta/\gamma} \left| \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^{[ns]} 1_{\{X_i^{(n)} > xU(\frac{n}{k})\}} - x^{-1/\gamma} C(s) \right) - \tilde{W}(x^{-1/\gamma}, C(s)) \right| \rightarrow 0 \text{ a.s.} \quad (5.6)$$

Proof Set $x_n := \frac{n}{k} (1 - F(xU(\frac{n}{k})))$. By the domain of attraction condition (1.5), we have $x_n \rightarrow x^{-1/\gamma}$, as $n \rightarrow \infty$, uniformly for all $x \geq x_0$. It easily follows from the proof that Theorem 5.4 remains true if we extend the domain of the STEP to $(t, s) \in (0, 2x_0^{-1/\gamma}] \times [0, 1]$. Therefore, we may replace t in (5.4) with x_n to obtain that

$$\sup_{0 \leq s \leq 1, x \geq x_0} x_n^{-\eta} \left| \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^{[ns]} 1_{\{X_i^{(n)} > xU(\frac{n}{k})\}} - x_n C(s) \right) - \tilde{W}(x_n, C(s)) \right| \rightarrow 0 \text{ a.s.} \quad (5.7)$$

The proof will be finished once we show that x_n can be replaced by its limit $x^{-1/\gamma}$ at the three places in this expression.

By (2.5) we obtain that (cf. de Haan and Ferreira (2006, p. 161)) for any $\delta > 0$ and sufficiently large n

$$\left| \frac{x_n - x^{-1/\gamma}}{A_2(n/k)} - x^{-1/\gamma} \frac{x^{\rho/\gamma} - 1}{\rho\gamma} \right| \leq \delta x^{(-1+\rho)/\gamma} \max(x^\delta, x^{-\delta}),$$

uniformly for all $x \geq x_0$. It follows that

$$\sup_{x \geq x_0} \left| \frac{x_n - x^{-1/\gamma}}{A_2(n/k) x^{-1/\gamma}} \right| = O(1), \quad n \rightarrow \infty.$$

Since $A_2(n/k) \rightarrow 0$, as $n \rightarrow \infty$, we may replace $x_n^{-\eta}$ with $x^{\eta/\gamma}$ in (5.7), and since $\sqrt{k}A_2(n/k) \rightarrow 0$, as $n \rightarrow \infty$, we may replace $x_n C(s)$ with $x^{-1/\gamma} C(s)$ in (5.7). The (uniform) continuity of the weighted bivariate Wiener process implies that, as $n \rightarrow \infty$,

$$\sup_{0 \leq s \leq 1, x \geq x_0} x^{\eta/\gamma} \left| \tilde{W}(x_n, C(s)) - \tilde{W}(x^{-1/\gamma}, C(s)) \right| \rightarrow 0. \quad \square$$

6 Proofs

Proof of Theorem 2.1 Taking $s = 1$ and $\eta = 0$ in (5.4), (with domain of t extended to $[0, 2]$) yields, as $n \rightarrow \infty$,

$$\sup_{0 \leq t \leq 2} \left| \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^n 1_{\{X_i^{(n)} > U(\frac{n}{kt})\}} - t \right) - \tilde{W}(t, 1) \right| \rightarrow 0 \text{ a.s.}$$

Applying Vervaat's lemma we obtain

$$\sup_{0 \leq t \leq 1} \left| \sqrt{k} \left(\frac{n}{k} (1 - F(X_{n, n-[kt]})) - t \right) + \tilde{W}(t, 1) \right| \rightarrow 0 \text{ a.s.}$$

Taking $t = 1$ and denoting $t_n := \frac{n}{k} (1 - F(X_{n, n-k}))$, we obtain that, as $n \rightarrow \infty$,

$$\left| \sqrt{k} (t_n - 1) + \tilde{W}(1, 1) \right| \rightarrow 0 \text{ a.s.} \quad (6.1)$$

We can thus replace t with t_n in (5.4) (with domain of t extended to $[0, 2]$) and obtain that

$$\sup_{0 \leq s \leq 1} \left| \sqrt{k} (\hat{C}(s) - t_n C(s)) - \tilde{W}(t_n, C(s)) \right| \rightarrow 0 \text{ a.s.} \quad (6.2)$$

By applying (6.1) to (6.2), together with the continuous sample path property of the Wiener process, we get that, as $n \rightarrow \infty$,

$$\sup_{0 \leq s \leq 1} \left| \sqrt{k} (\hat{C}(s) - C(s)) - (\tilde{W}(1, C(s)) - C(s) \tilde{W}(1, 1)) \right| \rightarrow 0 \text{ a.s.} \quad (6.3)$$

Defining the standard Brownian bridge $B(u) = \tilde{W}(1, u) - u \tilde{W}(1, 1)$ completes the proof of the first statement in the theorem.

Next, we prove the second statement, the asymptotic normality of the Hill estimator.

Taking $s = 1$ and $x_0 = \frac{1}{2}$ in (5.6) yields, as $n \rightarrow \infty$,

$$\sup_{x \geq \frac{1}{2}} x^{\eta/\gamma} \left| \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^n 1_{\{X_i^{(n)} > x U(\frac{n}{k})\}} - x^{-1/\gamma} \right) - \tilde{W}(x^{-1/\gamma}, 1) \right| \rightarrow 0 \text{ a.s.} \quad (6.4)$$

The limit relation (6.4) is the same as that for the tail empirical process based on i.i.d. observations, see de Haan and Ferreira (2006, Theorem 5.1.4). Therefore, the asymptotic normality of the Hill estimator, which can be proved via the tail empirical process, follows, see de Haan and Ferreira (2006, Example 5.1.5). More precisely, we obtain, as $n \rightarrow \infty$, that

$$\sqrt{k}(\hat{\gamma}_H - \gamma) \rightarrow \gamma \left(\int_0^1 \tilde{W}(t, 1) \frac{dt}{t} - \tilde{W}(1, 1) \right) \text{ a.s.}$$

It readily follows that $N_0 := \int_0^1 \tilde{W}(t, 1) \frac{dt}{t} - \tilde{W}(1, 1)$ is standard normal. Finally, it is easy to check that B and $\tilde{W}(\cdot, 1)$, and hence B and N_0 , are independent. \square

Proof of Theorem 3.1 From (5.6) we obtain, as $n \rightarrow \infty$,

$$\begin{aligned} & \sup_{0 \leq s_1 < s_2 \leq 1, s_2 - s_1 \geq \delta} \sup_{x \geq x_0} x^{\eta/\gamma} \left| \sqrt{k} \left(\frac{1}{k} \sum_{i=[ns_1]+1}^{[ns_2]} 1_{\{X_i^{(n)} > xU(\frac{n}{k})\}} - x^{-1/\gamma} (C(s_2) - C(s_1)) \right) \right. \\ & \quad \left. - \left(\tilde{W}(x^{-1/\gamma}, C(s_2)) - \tilde{W}(x^{-1/\gamma}, C(s_1)) \right) \right| \rightarrow 0 \text{ a.s.} \end{aligned} \quad (6.5)$$

From (6.3), we obtain that eventually for all s_1, s_2 such that $s_2 - s_1 \geq \delta$,

$$\hat{C}(s_2) - \hat{C}(s_1) > \frac{1}{2}(C(s_2) - C(s_1)) > \frac{1}{2}b\delta > 0 \text{ a.s.}$$

Hence, dividing (6.5) by $\hat{C}(s_2) - \hat{C}(s_1)$, yields, as $n \rightarrow \infty$,

$$\begin{aligned} & \sup_{0 \leq s_1 < s_2 \leq 1, s_2 - s_1 \geq \delta} \sup_{x \geq x_0} x^{\eta/\gamma} \left| \sqrt{k} \left(\frac{1}{k(s_1, s_2)} \sum_{i=[ns_1]+1}^{[ns_2]} 1_{\{X_i^{(n)} > xU(\frac{n}{k})\}} - x^{-1/\gamma} \frac{C(s_2) - C(s_1)}{\hat{C}(s_2) - \hat{C}(s_1)} \right) \right. \\ & \quad \left. - \frac{\tilde{W}(x^{-1/\gamma}, C(s_2)) - \tilde{W}(x^{-1/\gamma}, C(s_1))}{\hat{C}(s_2) - \hat{C}(s_1)} \right| \rightarrow 0 \text{ a.s.} \end{aligned} \quad (6.6)$$

Similarly we obtain from (6.3) that almost surely, as $n \rightarrow \infty$,

$$\sup_{0 \leq s_1 < s_2 \leq 1, s_2 - s_1 \geq \delta} \left| \sqrt{k} \left(\frac{\hat{C}(s_2) - \hat{C}(s_1)}{C(s_2) - C(s_1)} - 1 \right) - \left(\frac{\tilde{W}(1, C(s_2)) - \tilde{W}(1, C(s_1))}{C(s_2) - C(s_1)} - \tilde{W}(1, 1) \right) \right| \rightarrow 0.$$

Hence, we can replace $\hat{C}(s_2) - \hat{C}(s_1)$ by $C(s_2) - C(s_1)$ in (6.6) and obtain that

$$\begin{aligned} & \sup_{0 \leq s_1 < s_2 \leq 1, s_2 - s_1 \geq \delta} \sup_{x \geq x_0} x^{\eta/\gamma} \left| \sqrt{k} \left(\frac{1}{k(s_1, s_2)} \sum_{i=[ns_1]+1}^{[ns_2]} 1_{\{X_i^{(n)} > xU(\frac{n}{k})\}} - x^{-1/\gamma} \right) \right. \\ & \quad \left. - L(x^{-1/\gamma}, s_1, s_2) \right| \rightarrow 0 \text{ a.s.,} \end{aligned} \quad (6.7)$$

where

$$\begin{aligned} L(v, s_1, s_2) &:= \frac{\tilde{W}(v, C(s_2)) - \tilde{W}(v, C(s_1))}{C(s_2) - C(s_1)} \\ &\quad - v \left(\frac{\tilde{W}(1, C(s_2)) - \tilde{W}(1, C(s_1))}{C(s_2) - C(s_1)} - \tilde{W}(1, 1) \right). \end{aligned}$$

Observe that the limit relation (6.7) gives uniformly asymptotic properties of pseudo-tail empirical processes based on observations from subsamples satisfying $s_2 - s_1 \geq \delta$. It is comparable with the limit relation (5.1.18) in de Haan and Ferreira (2006), which is the basis for proving the asymptotic normality of the Hill estimator.

Next, we establish a uniform analog of the relation (5.1.19) therein. For notational convenience, set $\tilde{k} := k_{(s_1, s_2]}$ and $\tilde{n} := [ns_2] - [ns_1]$. Order the observations $X_{[ns_1]+1}, \dots, X_{[ns_2]}$ as $X_{s_1, s_2, 1} \leq \dots \leq X_{s_1, s_2, \tilde{n}}$. Taking $\eta = 0$ in (6.7) and applying a generalized Vervaat lemma, see Einmahl et al. (2010, Lemma 5), yields

$$\sup_{0 \leq s_1 < s_2 \leq 1, s_2 - s_1 \geq \delta} \sup_{\frac{1}{2} \leq t \leq 2} \left| \sqrt{k} \left(\frac{X_{s_1, s_2, \tilde{n} - [\tilde{k}t]}}{U(n/k)} - t^{-\gamma} \right) - \gamma t^{-\gamma-1} L(t, s_1, s_2) \right| \rightarrow 0 \text{ a.s.},$$

as $n \rightarrow \infty$. By taking $t = 1$, we obtain that, as $n \rightarrow \infty$,

$$\sup_{0 \leq s_1 < s_2 \leq 1, s_2 - s_1 \geq \delta} \left| \sqrt{k} \left(\frac{X_{s_1, s_2, \tilde{n} - \tilde{k}}}{U(n/k)} - 1 \right) - \gamma L(1, s_1, s_2) \right| \rightarrow 0 \text{ a.s.}, \quad (6.8)$$

which is a uniform analog of relation (5.1.19) in de Haan and Ferreira (2006). Using (6.7) and (6.8) in a similar way as in Example 5.1.5 therein, yields, as $n \rightarrow \infty$,

$$\sup_{0 \leq s_1 < s_2 \leq 1, s_2 - s_1 \geq \delta} \left| \sqrt{k} (\hat{\gamma}_{(s_1, s_2]} - \gamma) - \gamma \left(\int_0^1 L(u, s_1, s_2) \frac{du}{u} - L(1, s_1, s_2) \right) \right| \rightarrow 0 \text{ a.s.}$$

We have

$$\begin{aligned} & \int_0^1 L(u, s_1, s_2) \frac{du}{u} - L(1, s_1, s_2) \\ &= \frac{\int_0^1 \tilde{W}(u, C(s_2)) - \tilde{W}(u, C(s_1)) \frac{du}{u}}{C(s_2) - C(s_1)} - \frac{\tilde{W}(1, C(s_2)) - \tilde{W}(1, C(s_1))}{C(s_2) - C(s_1)} \\ &= \frac{\left(\int_0^1 \tilde{W}(u, C(s_2)) \frac{du}{u} - \tilde{W}(1, C(s_2)) \right) - \left(\int_0^1 \tilde{W}(u, C(s_1)) \frac{du}{u} - \tilde{W}(1, C(s_1)) \right)}{C(s_2) - C(s_1)}. \end{aligned}$$

The proof is completed by noting that the process W defined by

$$W(s) := \int_0^1 \tilde{W}(u, s) \frac{du}{u} - \tilde{W}(1, s),$$

is a standard Wiener process. □

Proof of Corollary 3.2 Combining Theorem 2.1 with Theorem 3.1, we obtain

$$\sup_{0 \leq s_1 < s_2 \leq 1, \hat{C}(s_2) - \hat{C}(s_1) \geq \delta} \left| \sqrt{k} \left(\frac{\hat{\gamma}_{(s_1, s_2]}}{\hat{\gamma}_H} - 1 \right) - \left(\frac{W(C(s_2)) - W(C(s_1))}{C(s_2) - C(s_1)} - W(1) \right) \right| \rightarrow 0 \text{ a.s.}$$

The asymptotic result for T_3 follows from this in conjunction with again Theorem 2.1 and the continuity of the sample paths of W .

Finally we consider T_4 . From Theorem 3.1, Theorem 2.1, and the continuity of the sample paths of W , we obtain

$$\sup_{1 \leq j \leq m} \left| \sqrt{k} \left(\hat{\gamma}_{(l_{j-1}, l_j]} - \gamma \right) - m\gamma \left(W \left(\frac{j}{m} \right) - W \left(\frac{j-1}{m} \right) \right) \right| \rightarrow 0 \text{ a.s.},$$

which implies that

$$\sup_{1 \leq j \leq m} \left| \sqrt{k} \left(\frac{\hat{\gamma}_{(l_{j-1}, l_j]}}{\hat{\gamma}_H} - 1 \right) - \left(m \left(W \left(\frac{j}{m} \right) - W \left(\frac{j-1}{m} \right) \right) - W(1) \right) \right| \rightarrow 0 \text{ a.s.}$$

The asymptotic result for T_4 thus follows. \square

Acknowledgement Research of Laurens de Haan partially supported by DEXTE – Development of Extremes in Time and Space, project EXPL/MAT-STA/0622/2013 and by national funds through the Fundação Nacional para a Ciência e Tecnologia, Portugal – FCT under the project PEst-OE/MAT/UI0006/2014.

References

- S. Coles. *An introduction to statistical modeling of extreme values*. Springer Verlag, 2001.
- A.C. Davison and R.L. Smith. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society B*, 52(3):393–442, 1990.
- L. de Haan and A. Ferreira. *Extreme value theory: an introduction*. Springer Verlag, 2006.
- L. de Haan and U. Stadtmüller. Generalized regular variation of second order. *Journal of Australian Mathematical Society, Series A*, 61:381–395, 1996.
- L. de Haan, A. Klein Tank, and C. Neves. On tail trend detection: modeling relative risk. *arXiv preprint arXiv:1106.4149*, 2011.
- H. Drees. Weighted approximations of tail processes for β -mixing random variables. *The Annals of Applied Probability*, 10:1274–1301, 2000.
- H. Drees and H. Rootzén. Limit theorems for empirical processes of cluster functionals. *The Annals of Statistics*, 38(4):2145–2186, 2010.
- J.H.J. Einmahl, M. Gantner, and G. Sawitzki. Asymptotics of the shorth plot. *Journal of Statistical Planning and Inference*, 140(11):3003–3012, 2010.
- P. Hall and N. Tajvidi. Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science*, 15(2):153–167, 2000.

- T. Hsing. On tail index estimation using dependent data. *The Annals of Statistics*, 19(3):1547–1569, 1991.
- D.W. Jansen and C.G. de Vries. On the frequency of large stock returns: Putting booms and busts into perspective. *The Review of Economics and Statistics*, 73(1):18–24, 1991.
- M.C. Jones. Simple boundary correction for kernel density estimation. *Statistics and Computing*, 3(3):135–146, 1993.
- P. Kearns and A. Pagan. Estimating the density tail index for financial time series. *The Review of Economics and Statistics*, 79(2):171–175, 1997.
- M.R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*, volume 11. Springer Verlag, 1983.
- D Mejlzer. On the problem of the limit distribution for the maximal term of a variational series. *L'vov Politechn Inst. Naučn. Zap. Ser. Fiz.-Mat.*, 38:90–109, 1956.
- C. Quintos, Z. Fan, and P.C.B. Phillips. Structural change tests in tail behaviour and the Asian crisis. *The Review of Economic Studies*, 68(3):633–663, 2001.
- G.R. Shorack and J.A. Wellner. *Empirical processes with applications to statistics*. Wiley, New York, 1986.
- A.W. van der Vaart and J.A. Wellner. *Weak convergence and empirical processes*. Springer, 1996.
- I. Weissman. Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815, 1978.

John H.J. Einmahl	Laurens de Haan	Chen Zhou
Dept. of Econometrics & OR	Department of Economics	Economic and Research Division
CentER, Tilburg University	Erasmus University	De Nederlandsche Bank
P.O. Box 90153	P.O. Box 1738	P.O. Box 98
5000 LE Tilburg	3000 DR Rotterdam	1000 AB Amsterdam
The Netherlands	The Netherlands	The Netherlands
j.h.j.einmahl@uvt.nl	ldehaan@ese.eur.nl	c.zhou@dnb.nl; zhou@ese.eur.nl